

*THE EFFECTS OF CONDUCTING BEHAVIORAL  
OBSERVATIONS ON THE BEHAVIOR OF  
THE OBSERVER*

ALICIA M. ALVERO

QUEENS COLLEGE, CITY UNIVERSITY OF NEW YORK

AND

JOHN AUSTIN

WESTERN MICHIGAN UNIVERSITY

Behavior-analytic approaches to occupational safety are often effective for improving safety in organizations, and have been successful in a wide variety of settings. The effects of these safety processes are thought to arise primarily from the behavioral observation process and the delivery of feedback. Typically, supervisors or employee observers involved in behavioral safety implementations conduct observations. The present study was an attempt to assess the effects of conducting observations on an observer's safety performance. An ABC multiple baseline counterbalanced across two sets of behaviors was conducted in a simulated office. The target behaviors involved knee and back positions during lifts; back, shoulder, and feet positions while sitting; neck and wrist positions while typing; and neck position during phone use. Substantial improvements in safety performance occurred after participants conducted observations on a videotape of a confederate's performance. The possible behavioral functions responsible for this change, and the implications of these findings for applied settings, are discussed.

DESCRIPTORS: behavioral safety, office safety behaviors, conducting observations, observer performance

---

In 2000 alone, a total of 5.3 million non-fatal injuries were reported in private industry workplaces; these injuries cost the country nearly \$21 billion per year (U.S. Bureau of Labor Statistics, 2001). Occupational accidents have contributed to growing workers' compensation premium payments, which totaled \$57.1 billion in 1995 (U.S. Bureau of Labor Statistics, 1999).

The behavior-analytic approach to occupational safety has proven to be effective in changing behavior and reducing injuries and

related costs (e.g., Fox, Hopkins, & Anger, 1987; Sulzer-Azaroff, Loafman, Merante, & Hlavacek, 1990) and has demonstrated success across various settings and populations (e.g., increasing safety of roofing crews, Austin, Kessler, Riccobono, & Bailey, 1996; increasing safety of industrial workers, Sulzer-Azaroff & de Santamaria, 1980; increasing safety of office personnel, McCann & Sulzer-Azaroff, 1996). Prior research has identified the principal components of an effective behavior-based safety process, which usually include assessment and identification of performance targets, development and implementation of a behavioral observation process, review of observation data, and implementation of behavioral feedback. The assessment involves a detailed review of an organization's injury records, first-aid logs, interviews with employees, and a review of any other records that may provide useful

---

The research presented in this article was part of the first author's thesis at Western Michigan University, which was funded, in part, by the Western Michigan University Graduate College Travel and Research Grant. We thank Terry McSween for his valuable input during the initial development of this research.

Address correspondence to Alicia M. Alvero, who is now at the Department of Psychology, Queens College, CUNY, Flushing, New York 11367 (e-mail: alicia.alvero@qc.edu).

data concerning previous safety efforts (Sulzer-Azaroff & Fellner, 1984). From this information, performance targets (or at-risk behaviors) are identified and compiled into an observation checklist.

The next step involves developing an observation process and training employees to conduct safety observations<sup>1</sup> using the checklist. When conducting observations, observers approach other employees and observe the behaviors listed on the checklist. Usually, the observed employee's performance is scored on the checklist, and the observer delivers immediate verbal feedback. The data from the checklists are then converted into safety percentages that are reviewed by either supervisors or employees or are graphed and posted in a prominent place for all employees to see. Frequent observation and reviews of the data help the safety professionals to adapt the process as needed.

The majority of previous research in this area has focused on the effects of feedback on organizational performance (see Alvero, Bucklin, & Austin, 2001, and Balcazar, Hopkins, & Suarez, 1985, for reviews of the feedback literature), but there are other independent variables in the safety process that have been less adequately studied. One of these is the observation process itself. In the behavior-analytic approach to occupational safety research literature, researchers, and occasionally supervisors, conduct observations. However, in employee-driven programs (i.e., in real-world application), most or all employees are trained to conduct behavioral observations. Interestingly, the employees who conduct the observations are the same employees who should engage in the safety tar-

get behaviors on the checklist when they are, in turn, observed. Although behavior analysts have closely examined and discussed the behavior of observers (e.g., Kazdin, 1977; Repp, Nieminen, Olinger, & Brusca, 1988), the focus of such discussion has been on experimental observers (i.e., outsiders to the population under study). Because most applied behavior analysis research describes programs implemented and evaluated by researchers (and not employees at the research site), questions focused on employee observers who are part of the internal workings of the system have not been widely studied. Therefore, the purpose of the present study was to examine the effects of the observation process, specifically to determine if the process of conducting safety observations of target behaviors at one point in time has an effect on the observer's subsequent performance of those same behaviors.

## METHOD

### *Participants, Setting, and Sessions*

Participants were 8 undergraduate university students (7 women and 1 man), and the study took place in simulated offices at a university campus. The offices were two observation rooms equipped with corner-mounted video cameras, which were used to record all sessions. For the purpose of this study, each observation room was furnished with the following equipment to resemble an office workstation: a desk, two chairs, a computer, a telephone, and an empty cardboard box. All observation sessions were 15 min in length.

### *Definition and Measurement of Dependent Variables*

All dependent variables, except lifting, were defined in terms of the percentage of intervals in which they occurred. Percentage occurrence was calculated by adding the number of intervals in which each safe be-

<sup>1</sup> When a person conducts a safety observation as part of a behavioral safety process, he or she must observe and evaluate another person's behavior and determine if it is safe or unsafe. Therefore, throughout this paper, "conducting observations" will refer to the observation and evaluation of another person's safety performance.

havior was observed, dividing it by the total number of observation intervals, and multiplying by 100%. Safe lifting frequency was counted and reported as a percentage of total number of lifts. Each task was comprised of several target behaviors; for example, performing a lift required engaging in two target behaviors: keeping the back aligned and bending at the knees. The eight target behaviors and their definitions were as follows: When lifting or putting down an object, (a) the back should be straight in a natural upright position throughout the lift, the back should not be twisted or parallel to the floor and (b) knees should be slightly bent (120° angle is recommended). When typing, (c) wrists should be in line with forearms and elbows, they should not be bent upwards or downwards, and (d) the neck should be upright and aligned with the back, and eyes should be level with the screen and document. When sitting, (e) the back should be upright and parallel against the back of the chair (not leaning at an angle against it), (f) shoulders should be aligned with the back and not slouched forward, and (g) both feet should be flat on the floor (ball of foot and heel should touch the floor). When using the telephone, (h) the neck should be upright and aligned with the back. Government ergonomic reports were reviewed to determine appropriate behaviors and their operational definitions (Occupational Safety and Health Administration, 1998; Office of Health and Safety Information System, 1998).

Each session was videotaped and scored at a later time by a trained observer using a checklist that contained each of the target behaviors and the definitions listed above. A momentary time-sampling procedure was used for data collection. Every 30 s, data were collected for behaviors occurring at that moment, with the exception of lifting. Target behaviors were scored in a specified order. For example, if a participant was using

the phone, the coder scored the four relevant target behaviors that occurred at the 30-s interval, in a specific order: back, shoulders, and feet (for sitting); and neck (while using the phone). Because lifting occurred so infrequently, the chances were very high that it would not be detected using the time-sampling procedure. Therefore, data were collected for every lift, regardless of when it occurred. A behavior was scored as being safe when it satisfied the definition listed above.

#### *Interobserver Agreement*

A second observer independently scored 40% of the sessions. An agreement was defined as an occurrence in which both observers scored the same mark (safe or unsafe) for each particular behavior. Interobserver agreement was calculated as the number of agreements divided by the number of agreements plus disagreements multiplied by 100%. Overall agreement between observers averaged 96.9% (range, 89% to 100%).

#### *Procedure*

*Baseline.* Before the start of the study, participants were randomly assigned to Group A or B. At the start of each baseline session, all participants were handed a list of instructions and were asked to perform the tasks described in the instructions. The tasks were (a) typing a few paragraphs using a word processor on a computer; (b) dialing a phone number and leaving a brief message on an answering machine; and (c) picking up a cardboard box containing five pieces of paper, placing it on a chair, removing a specific piece of paper, and placing the box back down on the floor. Each task was repeated at least four times, and the order of tasks varied from session to session.

*Information phase.* During the first session of the information phase, participants were informed about the nature of the study. All participants were told that the purpose of the study was to observe individual safety

behaviors in an office setting and were then given a handout containing definitions for four of the eight target behaviors and how to perform them safely. Group A received information on one set of four behaviors (back straight and knees bent when lifting, neck and wrist position while typing), and Group B received information on the other four behaviors (back, shoulder, and feet position when sitting; neck alignment when using the phone). Participants were given this information at the start of each session and were asked to study it for 5 min before they were handed the list of tasks to perform. The remainder of the session followed the same procedures as those during baseline. The purpose of this intervention phase was twofold: to eliminate demand characteristics that are often displayed by participants taking part in a laboratory study (Kazdin, 1992) and to separate the effects of information from those of conducting safety observations. Often participants in a laboratory study try to please the experimenter by engaging in the behaviors that they believe the experimenter is interested in observing; this is referred to as a demand characteristic. The information phase was created to minimize the effects of demand characteristics during the observation phase.

*Observation phase.* Before being given a list of tasks to perform, participants were asked to observe a 5-min videotape of an experimental confederate performing tasks similar to those the participant performed during each session. Participants were then asked to collect data on the confederate's safety performance using a checklist containing the same four target behaviors they had been given during the information phase (Group A scored the first four; Group B scored the second four) and definitions on how to perform each safely. Participants scored the videotape using an event-recording procedure; thus, a behavior was scored immediately after observing its occurrence.

Participants scored a different 5-min videotape before every session during the observation phase, and all videotapes included a random combination of safe and unsafe behaviors. Immediately after scoring the videotape, participants were instructed to perform the tasks as in the baseline condition.

### *Experimental Design*

The experimental design included components of a multielement and a multiple baseline design. Initially, a multielement design was used to compare the two interventions (information and observation) with baseline. As indicated above, participants were randomly assigned to Group A or Group B. In accordance with a multielement design, the first four target behaviors were sequentially exposed to the information and observation conditions for Group A while the remaining four target behaviors remained in baseline. For Group B, the last four target behaviors were sequentially exposed to the information and observation conditions while the first four target behaviors remained in baseline. In accordance with a multiple baseline design, once stable levels of target behaviors were obtained in the observation condition for half of the target behaviors (i.e., the first four for Group A or the last four for Group B), the observation condition was introduced for the remaining target behaviors (i.e., the last four for Group A and the first four for Group B).

### *Integrity of the Independent Variables*

Scripts were developed and used for all of the verbal instructions that were given to the participants. This ensured that all participants were exposed to the same instructional set. During the information phase, participants were required to check off each behavior and safety definition on the information sheet and sign the sheet to ensure that they had been provided with information during each session throughout this

phase. The safety checklists used by the participants during the observation phase were collected to provide verification that participants conducted observations during the session. The videotapes shown to the participants were arranged in a specific order to ensure that all participants were exposed to the same sequence.

## RESULTS

Figures 1 and 2 show the safety performance of the 4 participants in Group A, and Figures 3 and 4 show results for the 4 participants in Group B. Each figure has two panels per participant, one showing the results for the first four target behaviors (e.g., Panel 1 in Figure 1 for Participant 1) and the other showing the results for the last four target behaviors (e.g., Panel 2 in Figure 1 for Participant 1).

Across the 8 participants, there were 32 opportunities for replication of target-behavior changes (four different behaviors for each of 8 participants) during the information phase. Safety performance remained unchanged from baseline for 25 of 32 opportunities (78%). Temporary increases were observed on three occasions (9%) for two different behaviors: lifting with knees (Participants 2 and 3) and neck position while typing (Participant 4). These increases were followed by gradual decreases toward baseline performance. Strong effects occurred for one behavior (sitting with feet on the floor) for Participants 5, 6, 7, and 8 (13%).

Across the 8 participants, there were 64 opportunities for replication of target-behavior changes (eight different behaviors for each of 8 participants) during the observation phase. No effects were observed for 4 of 64 opportunities (6%). Of these four opportunities, two failed to change because of a procedural shortcoming. Performance did not change for Participant 5 (Figure 3) on the two behaviors involved with lifting: back

alignment and knee bend. The reason no improvements occurred in this case was that the participant never left her chair to lift, and therefore she could not lift properly. We never instructed her to stand when lifting, because that would have amounted to changing the experimental protocol in a single case. No other participants acted in this manner, nor did they receive special instruction.

Gradual improvements in performance were observed for 18 of 64 opportunities (28%); these occurred most frequently in typing and lifting tasks. A clear example of a gradual improvement can be seen in the lifting performance of Participant 1 (Figure 1). Back alignment and correct knee bends increased gradually across each session during the observation phase. Both behaviors remained at 0% safe during the first session of the observation phase and gradually improved to 83% (back) and 100% (knees) safe during the last session of the phase.

Of the 64 behaviors targeted across 8 participants, 42 (66%) increased substantially after the implementation of the observation phase. The most dramatic of these changes can be seen in the lifting performance of Participant 6 (Figure 3). Both back and knee behaviors averaged 0% safe during baseline and 100% safe during the observation phase.

## DISCUSSION

The results of the current study suggest that improvements in performance occurred as a result of conducting observations of confederate performance. Using a checklist to observe and evaluate the safety performance of others increased the safety behavior of the observer. It is important to note that the increases in safety performance occurred as a result of observing and evaluating a confederate's safety performance; observing alone did not change safety performance. The multiple baseline design used in this

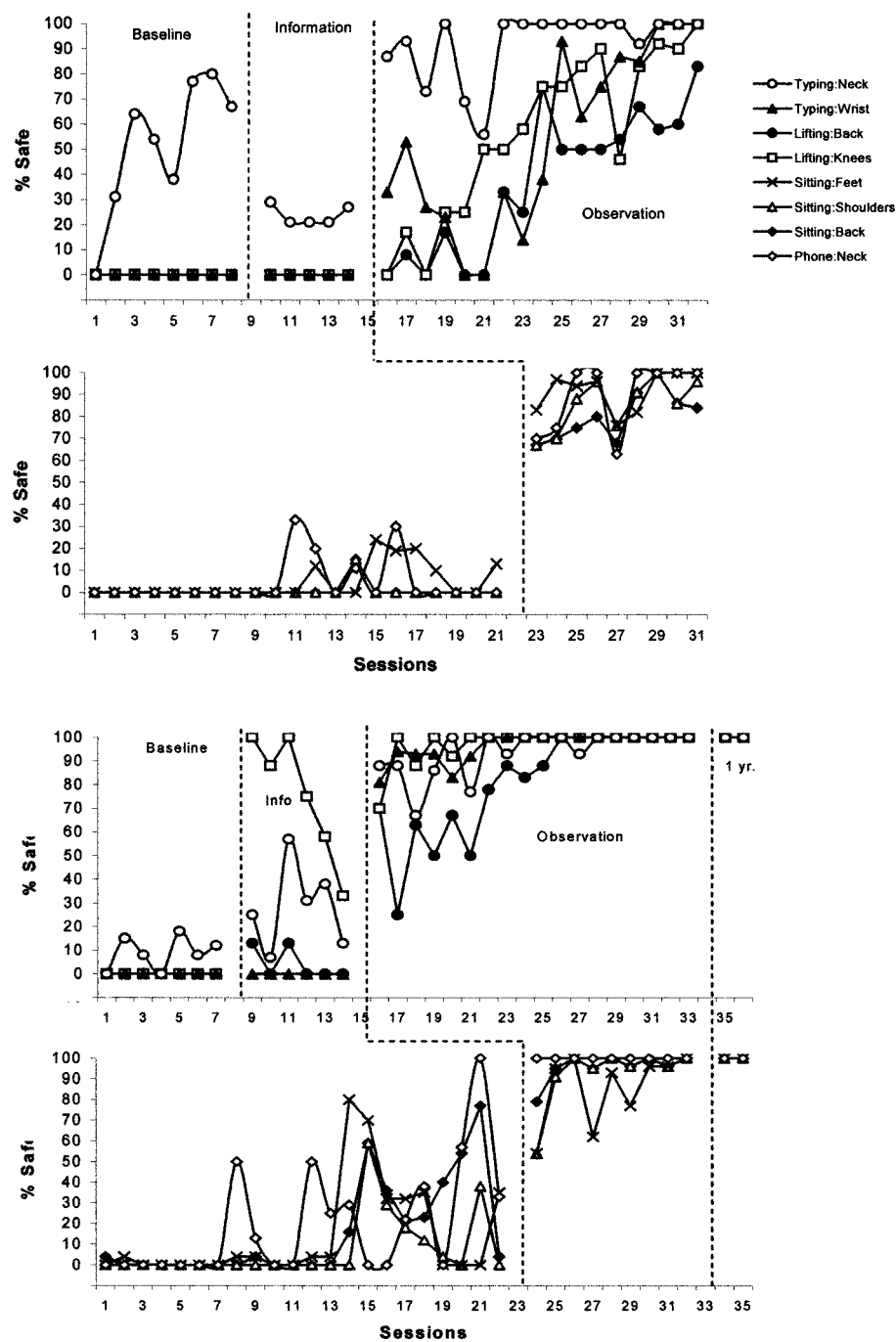


Figure 1. The percentages of target responses completed correctly during baseline, the information condition, and the observation condition for Participant 1 (top two panels) and Participant 2 (bottom two panels).

study allows us to draw this conclusion. Increases in safety performance occurred only on the behaviors targeted on the safety checklist. In other words, all eight target be-

haviors were observed when a participant scored a videotape, but increases occurred only on the behaviors being observed and evaluated using the checklist. It is unclear,

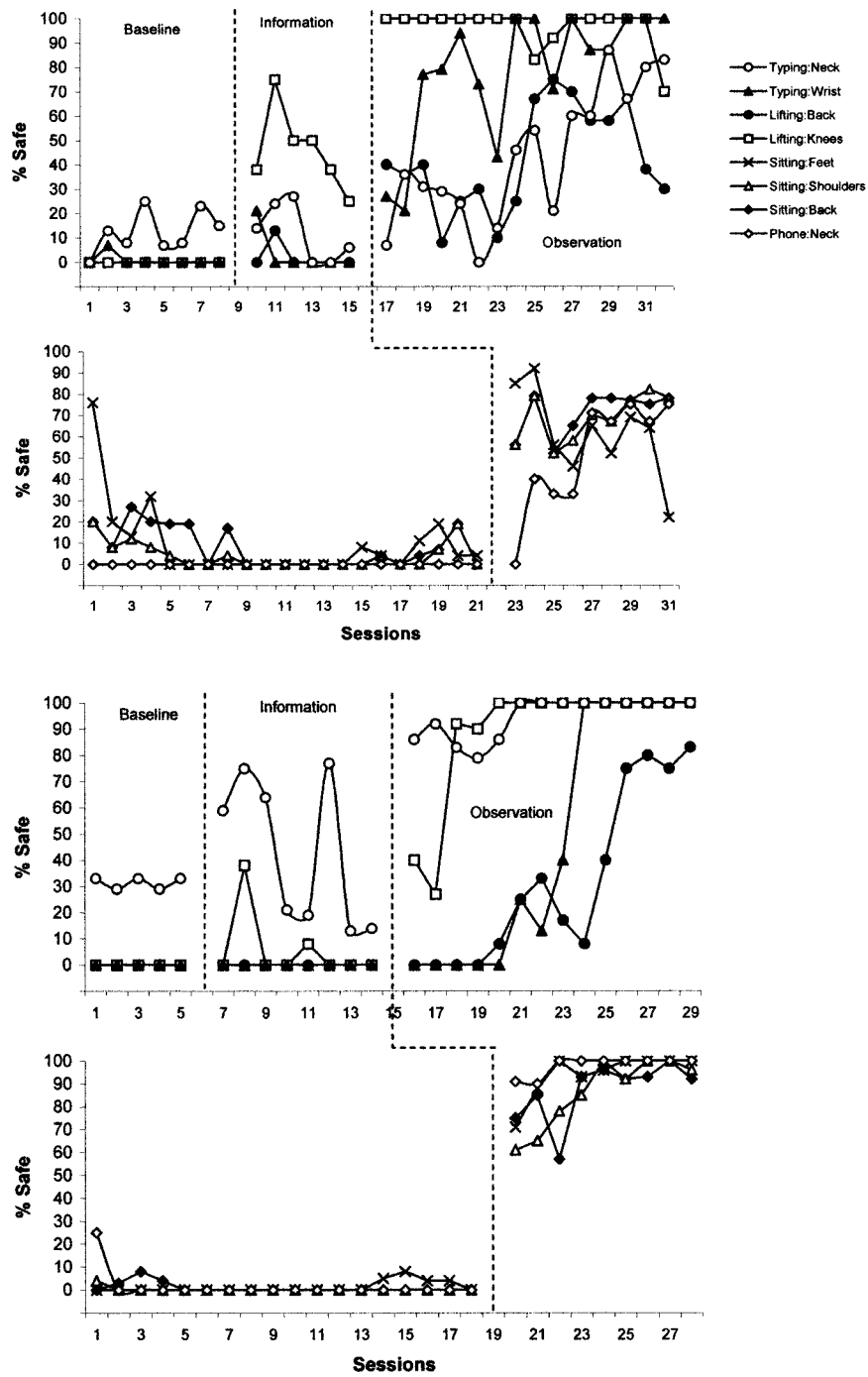


Figure 2. The percentages of target responses completed correctly during baseline, the information condition, and the observation condition for Participant 3 (top two panels) and Participant 4 (bottom two panels).

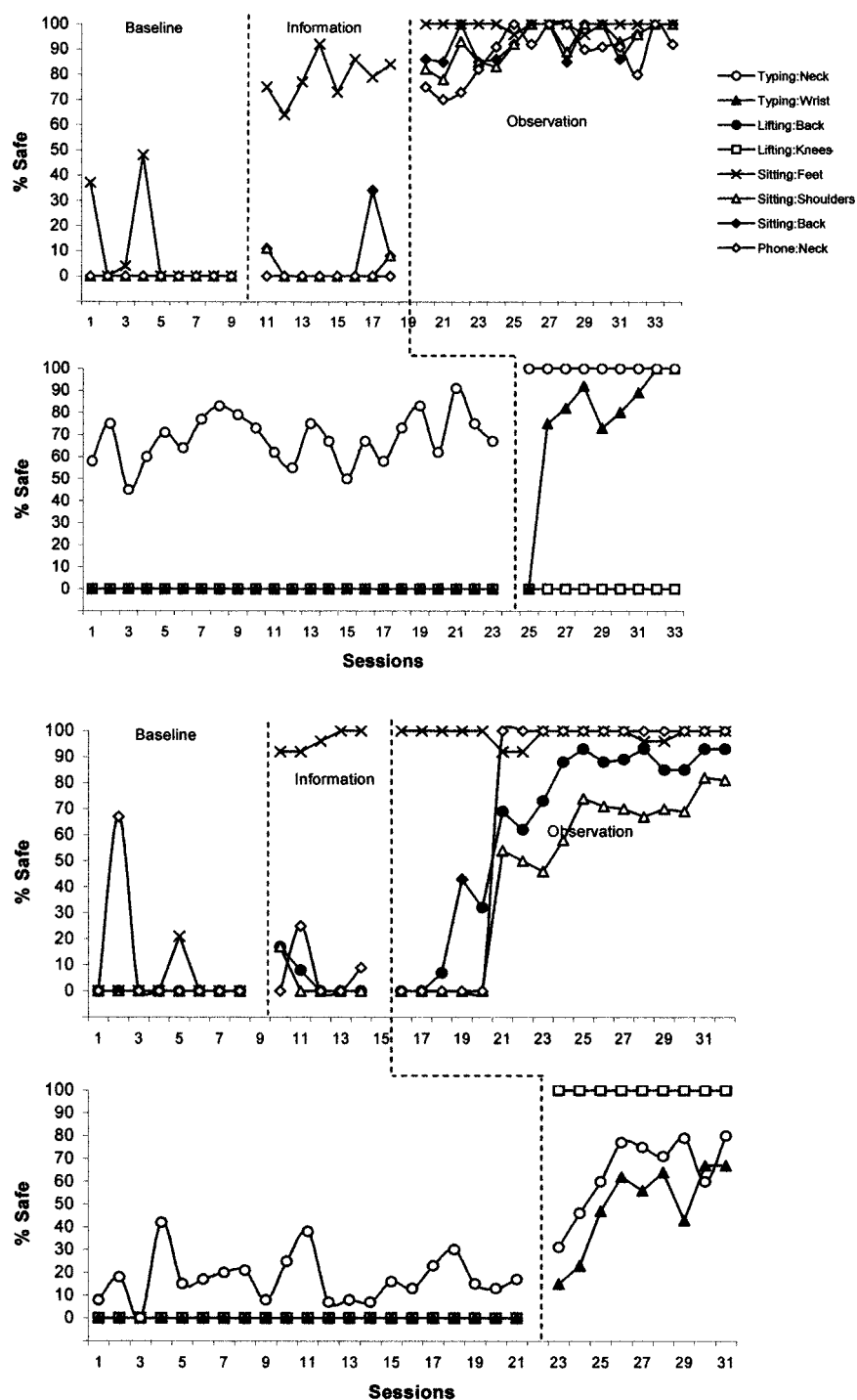


Figure 3. The percentages of target responses completed correctly during baseline, the information condition, and the observation condition for Participant 5 (top two panels) and Participant 6 (bottom two panels).

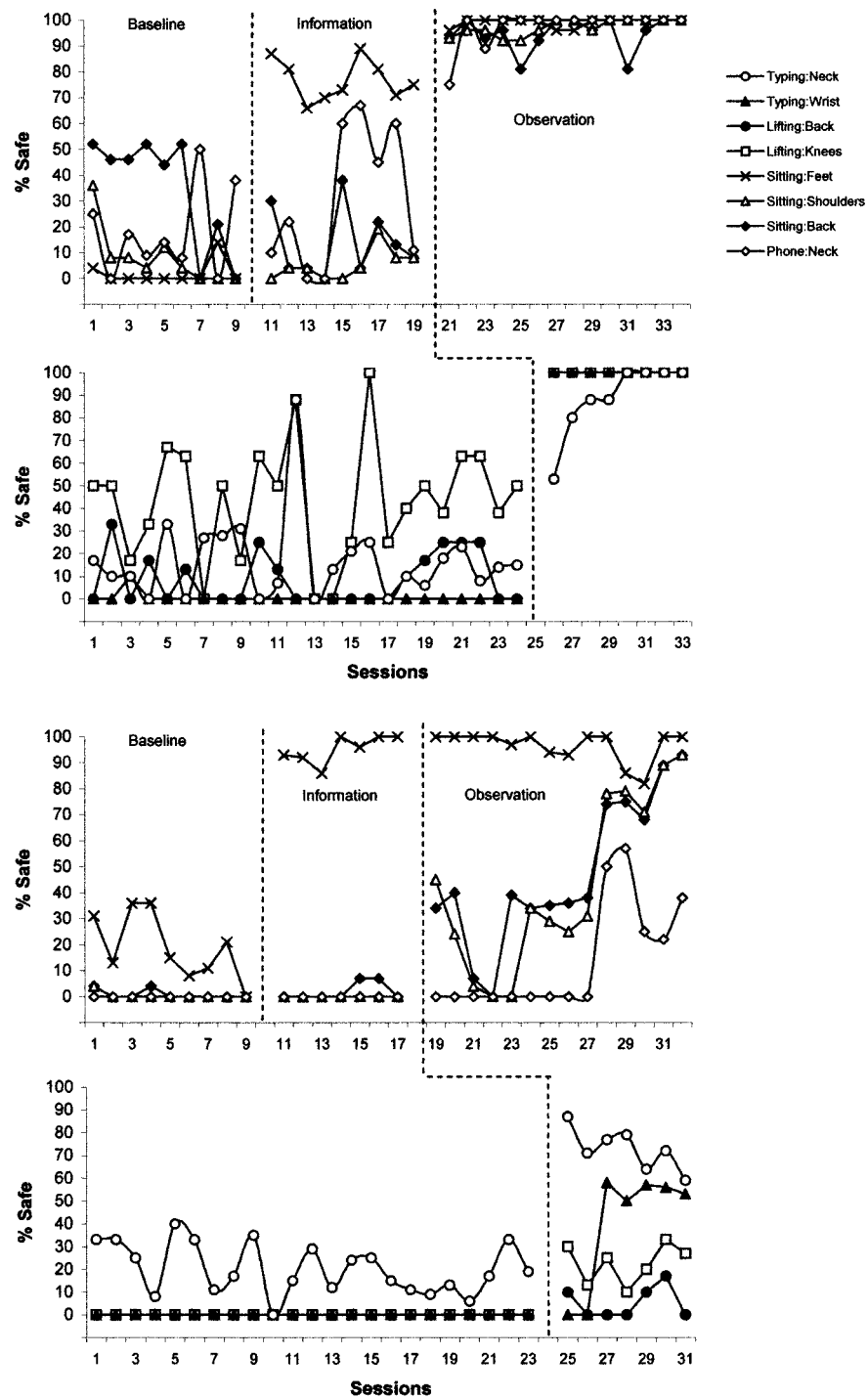


Figure 4. The percentages of target responses completed correctly during baseline, the information condition, and the observation condition for Participant 7 (top two panels) and Participant 8 (bottom two panels).

from a behavioral perspective, why these increases in the target behaviors occurred during the observation phase, and why they usually did not occur during the information phase. Another point of interest is why some behaviors increased immediately from 0% to 100%, whereas others showed more gradual increases.

Based on exit interviews with participants, we suspect that self-monitoring was a potential causal factor involved in the changes that occurred during the observation phase. Participants almost uniformly reported that scoring the videotape caused them to begin to self-monitor during their own sessions.

The current results add to the literature on safety performance in several ways. First, our literature search revealed no other attempts to test systematically the effects of conducting safety observations on observer behavior in the field of occupational safety. In addition, the effects of observing were robust, suggesting that this may be a simple and efficient method of increasing the safety performance of employees.

Performance data were collected for each participant and provided an important strength of the current study. Although future research is required to determine the behavioral functions of conducting observations, this study suggests that improvement in ergonomic safety does not negatively affect productivity. As another strength, the exit interviews provided important information concerning the possible causes of behavior change. We are currently conducting studies to investigate empirically these potential sources of control.

The laboratory setting used for the current research created both strengths and weaknesses. Participants were college students rather than actual employees. However, although the participants created no organizationally relevant products, it closely simulated jobs performed regularly by office personnel. A weakness of the study was that

we failed to perform complete independent-variable integrity checks for information and observation. It was verified that participants were exposed to the appropriate form of information and videotapes, and that participants scored the videotapes as planned. However, we did not test the knowledge of the participants after information exposure, nor did we measure the accuracy of observations during scoring. Future studies should attend to this deficiency accordingly.

Although the current research suggests that observing the behavior of others changes the behavior of the observer, future research should build on this experiment by manipulating variables related to the effect. For example, future studies could examine the potential moderating influence of observation accuracy, type and specificity of the behavioral checklist, and variation in properties of the responses observed.

If an observer effect exists in some or most cases of behavior-analytic approaches to occupational safety, practitioners may want to adjust methods of behavioral safety implementations to have all employees conduct observations. This could change considerably the focus of consultant-based behavioral safety implementations. Given a limited number of training days from consultants, safety managers may decide to devote more time to employee observation training than is currently the case.

## REFERENCES

- Alvero, A. M., Bucklin, B. R., & Austin, J. (2001). An objective review of the effectiveness and essential characteristics of performance feedback in organizational settings (1985–1998). *Journal of Organizational Behavior Management*, 21, 3–29.
- Austin, J., Kessler, M. L., Riccobono, J. E., & Bailey, J. S. (1996). Using feedback and reinforcement to improve the performance and safety of a roofing crew. *Journal of Organizational Behavior Management*, 16(2), 49–75.
- Balcazar, F., Hopkins, B. L., & Suarez, Y. (1985). A critical, objective review of performance feedback.

- Journal of Organizational Behavior Management*, 7(3-4), 65-89.
- Fox, D. K., Hopkins, B. L., & Anger, W. K. (1987). The long-term effects of a token economy on safety performance in open-pit mining. *Journal of Applied Behavior Analysis*, 20, 215-224.
- Kazdin, A. E. (1977). Artifact, bias, and complexity of assessment: The ABCs of reliability. *Journal of Applied Behavior Analysis*, 10, 141-150.
- Kazdin, A. E. (1992). *Methodological issues and strategies in clinical research*. Washington, DC: American Psychological Association.
- McCann, K. B., & Sulzer-Azaroff, B. (1996). Cumulative trauma disorders: Behavioral injury prevention at work. *Journal of Applied Behavioral Science*, 32, 227-291.
- Office of Health and Safety Information System. (1998). *Safety manuals*. Retrieved January 8, 1999, from <http://www.cdc.gov/od/ohs/default.htm>
- Occupational Safety and Health Administration. (1998). *OSHA ergonomics report DT931018*. Retrieved December 19, 1999, from <http://www.osha.gov>
- Repp, A. C., Nieminen, G. S., Olinger, E., & Brusca, R. (1988). Direct observation: Factors affecting the accuracy of observers. *Exceptional Children*, 55, 29-36.
- Sulzer-Azaroff, B., & de Santamaria, C. (1980). Industrial safety hazard reduction through performance feedback. *Journal of Applied Behavior Analysis*, 13, 287-295.
- Sulzer-Azaroff, B., & Fellner, D. (1984). Searching for performance targets in the behavioral analysis of occupational health and safety: An assessment strategy. *Journal of Organizational Behavior Management*, 6(2), 53-65.
- Sulzer-Azaroff, B., Loafman, B., Merante, R. J., & Hlavacek, A. C. (1990). Improving occupational safety in a large industrial plant: A systematic replication. *Journal of Organizational Behavior Management*, 11(1), 99-120.
- U.S. Bureau of Labor Statistics. (1999). *Statistical abstract of the United States: 1998*. Washington, DC: U.S. Bureau of the Census.
- U.S. Bureau Labor of Statistics. (2001). *United States Department of Labor news*. Retrieved December 21, 2001, from <http://www.bls.gov/iif/home.htm>

Received May 29, 2002

Final acceptance August 23, 2004

Action Editor, Mark O'Reilly

## STUDY QUESTIONS

1. Briefly describe the four principal components of an effective behavior-based safety process outlined by the authors.
2. Which aspect of behavior-based safety processes has been the typical focus of previous research, and which aspect was examined in this investigation?
3. Briefly describe the general experimental context and target behaviors.
4. How were the target behaviors measured?
5. Briefly describe each of the experimental conditions.
6. Summarize the results obtained during the information and observation conditions of the study.

7. According to the authors, what behavioral process probably accounted for the observed changes in performance, and what evidence did they offer to support the influence of this process?
8. The authors suggested that improvement in ergonomic safety does not affect productivity negatively. What data would support or refute this claim?

Questions prepared by Jennifer N. Fritz and Sarah E. Bloom, University of Florida